CrossMark

Looking for best performers: a pilot study towards the evaluation of science parks

M. Ferrara[1,3] · F. Lamperti[2,3,4] · R. Mavilia[3,4]

Abstract Science Parks are complex institutions that aim at promoting innovation and entrepreneurship at local level. Their activities entertain a large set of stakeholders going from internal and external researchers to entrepreneurs, local level public administration and universities. As a consequence, their performances extends on a large set of dimensions affecting each other. This feature makes Science Parks particularly difficult to be properly compared. However, evaluating their performances in a comparable way may be important for at least three reasons: (1) to identify best practices in each activity and allow a faster diffusion of these practices, (2) to inform potential entrepreneurs about institutions better supporting start-ups birth and first stages and (3) to guide public policies in the distribution of funds and incentives. The multidimensional nature of Science Parks raises the problem of aggregating performances in simple indexes that can be accessed by stakeholders willing to compare different structures on the basis of their own preferences. This paper exploits a new dataset on Italian Science Parks to provide a pilot study towards this direction. In particular, we apply Choquet integral based Multi-Attribute Value Theory to elicit stakeholders' preferences on different dimensions of Science Parks' performances and construct a robust index allowing to rank them. This tool can be used to support the decision making process of multiple stakeholders looking for best (or worst) performers and allows to account both for subjective nature of the evaluation process and the interactions among decision attributes. Despite the present study employs only a limited

✉ M. Ferrara
massimiliano.ferrara@unirc.it

1   Mediterranean University of Reggio Calabria, Reggio Calabria, Italy

2   Sant'Anna School of Advanced Studies, Pisa, Italy

3   CRIOS Bocconi University, Milan, Italy

4   MEDAlics, University for Foreigners "Dante Alighieri" of Reggio Calabria, Reggio Calabria, Italy

number of respondents and performance measures, the procedure we present can be straightforwardly adapted to much richer environments.

Introduction

In recent years a good number of countries has invested in the creation of a variety of bodies aimed at fostering an entrepreneurial culture based on innovation and research, and supporting already existing innovative start-ups. Among such bodies we find Science and Technology Parks, Research Parks, Incubators and many other typologies of institutions among which it is extremely difficult to distinguish with reasonable precision (Saublens et al. 2007; Link and Link 2003; Link and Scott 2003). In this paper we proposes an approach to evaluate and compare all structures that, at least, share both a research/innovation and an entrepreneurship oriented side. We refer to these structures, for sake of simplicity, as *Science Parks* (SPs). The complex set of activities SPs typically engage in, ranging from applied research to start-ups incubation, calls for a cautious approach to evaluate their performances (Luger and Goldstein 1991; Monck and Peters 2009). Comparing performances of different SPs is fundamental for a number of reasons: to identify best practices in each activity and allow a faster diffusion of these practices, to inform potential entrepreneurs about institutions better supporting start-ups birth and their first stages of life and, finally, to guide public policies in the distribution of funds and incentives. In addition, established companies might be interested as well in comparing performances of different SPs as this could influence their decision about where to locate research units. However, very few attempts have been undertaken to address the issue of evaluating SPs and none of them provide a tool that accounts for the endogenous preferences of multiple stakeholders. The present paper aims at filling this gap. In particular, we propose a methodology to construct an index that summarizes and aggregates multidimensional performances of each Science Park by mean of Choquet integration. There are two great advantages of this approach. First, it allows to take into account and analyse possible *complementarities* and *redundancies* among attributes characterizing performances. This is relevant when the aggregate performance level of a single structure might not be deduced from the simple averaging of single performances. For example, consider a Science Park whose incubated firms grow fast both in terms of revenues and employees. This means that they are both creating job opportunities and gaining competitiveness in the respective industries. A policy maker might judge these two features as relevant in assessing the overall performance of the park but, on the other side, it might as well recognize that one is typically correlated with the other (that is, firms that tend to growth in terms of sales or turnover, are prone to increase in size as well, and vice-versa, Delmar 2006). In such a situation every weighted average would fail to recognize that, possibly, these two dimensions of performance are redundant. Once one of the two is given high importance, there is no reason to do the same with the other but, on the same ground, assigning little importance to both will be conceptually erroneous as well. The approach we propose here, straightforwardly allows to include features of complementarity and

redundancy and quantify them for each pair of performance indicators. This leads to the second advantage provided by our method: it elicits preferences from a number of subjects and deduces the importance to be assigned to each dimension accordingly. Given that the extra-value brought by a composite indicator of performance lies precisely in its capacity to aggregate various dimensions into a single one in a non-arbitrary way, the set of weights used in its construction must be based on nothing else than individual preferences. Note that the elicitation of preferences provides a real advantage here. Indeed, when attributes of performance are numerous and possibly hierarchically structured, it might be difficult for a subject to express explicitly a set of weights. Our approach allows to deduce such measures from a series of simple rankings over SPs with different performances that subjects are asked to provide. In addition, the heterogeneity of preferences (each stakeholder possesses her own) can be preserved and analysed through a cluster analysis identifying groups of homogeneous responders. Subjects are represented as points in a suitable multidimensional space, whose centroid is then used as a synthetic indicator of aggregate preferences[1], which allows to compute a final unidimensional index of performance. Summarizing, we propose a tool that allows stakeholders to compare SPs and look for best (or worst) performers according to their own elicited preferences and accounting for multiple and possibly *interacting* dimensions. Interactions can also be quantified and their robustness analysed across the spectrum of stakeholders involved. Our tool is then applied to a pilot study consisting in the comparative evaluation of Italian Science Parks. Within this study, each structure is characterized by 8 dimensions of performance organized hierarchically. Stakeholders are loosely represented by a sample of entrepreneurship (master level) students and academic researchers. In particular, the relative importance of each dimension and its interaction with the others is attributed through a previous elicitation of preferences from 30 subjects including 10 academic researchers and 20 students from two Master of Science programs in Entrepreneurship and Innovation Management respectively. Despite this constitute only a first attempt to show how our tool might work, the same approach can easily adapted to much richer environments where both the sets of stakeholders and performance measures are enlarged and tailored on the specific decision problem.

The paper is organized as follows. "Literature review" section presents the emergence of attempts to evaluate SPs, especially handling the trade-off between research performance and success of on-park firms. "Methodology" section presents the main theoretical basis of the approach we use. "A pilot study: application to Italian Science Parks" section introduces our pilot study and discusses the application of our tool to the case of Italian SPs; "Results I: behavioural analysis of the Aggregation" section shows all information and results that our method can provide. Finally, "Results II: comparative analysis of Italian Science Parks" section concludes the paper.

Literature review

Science parks are nowadays largely regarded as key elements of the research based regional development policy (Saublens et al. 2007). However, evaluating their performances is a complex task; first because we lack a shared and clear taxonomy, which distinguishes between science parks and different structures. Secondly, there are poorly

---

[1] It is relevant to notice that different points other than the centroid can be used as *summary* for the population of subjects whose preferences has been elicited; however, our choice guarantees some desirable properties (see "Properties of the aggregation" section).

specified tools able to capture different dimensions of performance, and data availability is often limited (Guy 1996).

As Monck and Peters (2009) suggest, there are a number of reasons why performance assessments of Science Parks are important and highly relevant for the involved stakeholders. First of all, these structures are often financially supported by public sector bodies who expect evidence of the efficiency and effectiveness of their expenditure in order to decide how to allocate funds and resources. The identification of best practices can, in turn, enhance competitiveness of the overall regional and national innovation system, while the identification of best structures can give information to potential partners and workers. Finally, performance assessment is, as in the case of nearly all other profit oriented businesses, essential for managers and stakeholders to develop the science park model and/ or objectives and to rectify any shortcomings.

The majority of studies assessing performances of SPs focus on the their role in supporting on-park firms (Felsenstein 1994; Colombo and Delmastro 2002; Siegel et al. 2003, among the first contributions see) or regional development, which is usually measured through the number of job positions created (Luger and Goldstein 1991) or the regional GDP(Ferrara et al. 2014). All these studies isolate the SPs' effect on a single dimension of firms' performance using others as controls. Ferguson and Olofsson (2004) show there is no significant difference between on and off park firms with respect to employment and sales growth; Lamperti et al. (2015) focus on the relationships among R&D investments, innovativeness and firms' growth finding that the presence of research centres in the park and the number of linkages with universities induce larger investment in R&D and foster tenants' innovativeness, while growth remains a largely unexplained phenomenon. Löfsten and Lindelöf (2002) report that relationships with universities help sustain firms' growth, which is high in on-park firms. The innovativeness of parks' tenants is shown to outperform those of comparable out-of-park firms in Squicciarini (2009) with the driver being co-location and hence the number of firms SPs might host. Link and Scott (2003) indicate that research parks have a positive impact on universities' growth and profile. They enable universities to increase the number of publications, patents, facilitate transfer of technologies and easily place graduates. Recently, Minguillo et al. (2015) and Minguillo and Thelwall (2015) have investigated the role of SPs in the UK[2] in fostering scientific activity and cooperation through the analysis of scientific publications. Parks seem to have a positive impact on the overall level of collaboration and production of science and technology, suggesting that networking activities of SPs might be considered as elements of success. All this literature (which is only partially surveyed here) has contributed to the identification of different key elements to evaluate SPs, as they are those driving tenants superior performance with respect to matched samples.

Another branch of contributions, instead, tries to identify a set of goals, dimensions of performance and indicators which might be used to assess SPs' success or weaknesses. The ANGLE Technology Study (ANGLE 2003), commissioned by the UK Science Park Association, breaks down parks' performance into two categories: the economic performance of affiliated firms (measured by employment growth, turnover growth and access to finance) and their innovation and technology commercialisation performance (assessed through new products/services launched, patent applications, R&D investments and proportion of qualified scientists or engineers in the workforce). Monck (2010) identified instead a more refined taxonomy distinguishing between key performance indicators, intermediate output and short

---

[2] Note that they carefully distinguish between Science Parks and different entities as Research Parks , Technology Parks, Incubators and other organizations, while we insert all structures sharing both research activities and incubation ones within our broad categorization.

term management indicators of performance. Building on it, Dabrowska (2011) collects results from a workshop on SPs' performance measurement where a number of different stakeholders (from universities' representatives to on-park firms' founders and commercial investors) have been involved and surveyed. The output is a complex, numerous (more than 40) and heterogeneous set of indicators, which has emerged to differ strongly, passing from one group of stakeholders to another. What remains completely unclear and un-investigated, at least to the best of our knowledge, is the issue concerning how to aggregate these dimensions of performance and compare different SPs along both single performance indicators and in an absolute way. This paper propose an attempt to fill this gap in the literature by providing the advantage of a non-arbitrary aggregation which starts from preferences of a number of stakeholders belonging to different groups (two in our case, but the approach can be easily extended). Another advantage of this approach is that the elicitation procedure can be conducted very quickly through a questionnaire (30–45 min), resulting in a practical alternative that can be used during stakeholders' workshops.[3]

It should be noticed that in a previous study Ferrara and Mavilia (2012) proposed a first attempt to aggregate multidimensional performances of Science Parks through a simple weighting average of scores along single attributes, with weights assigned arbitrarily. We find relevant to recall that this approach is flawed by a series of problems. First, it does not explicitly accounts for individual's preferences in a rigorous way since researchers agreed on the weight without directly and completely expressing their preferences. Secondly, the aggregation approach should give the possibility to use a panel of experts, possibly with heterogeneous background and experience, to assign importance of different dimensions. Third, different people may not agree on the relative importance of each attribute. Accounting in the model for this heterogeneity in preferences could yield more reliable information than forcing people to agree on weights. Fourth, a simple weighted sum does not account for interactions between attributes, which might play a relevant role in expressing decision makers' preferences. Therefore, groups of attributes can count more (or less) than the simple sum of the relative weight of each component. The procedure outlined in the next sections allows to solve or, at least, to mitigate all these problems.

Methodology

We propose to use MAVT to extract, from empirical basis, plausible weights to be assigned to various dimensions of performances of Science Parks. As outlined by Meyer and Ponthière (2011), the extra-value brought by a composite indicator of performance lies precisely in its capacity to aggregate various dimensions into a single one in a non-arbitrary way. Hence, the set of weights used in its construction must be based on nothing else than individual preferences. Actually, the proposed methodology starts from simple orderings of Science Parks characterized by different performances and derive a numerical representation of the underlying preferences, where the weights assigned to the various attributes of SPs reflect the intensity of experts' subjective concern for those attributes.

While it is tempting, for simplicity, to represent preferences over multiattribute societies by means of a classical weighted sum (as, for example, in the Human Development

---

[3] Note that more precise alternatives do exist (Labreuche and Grabisch 2003, see for example the use of MACBETH in) Nevertheless, this task is not trivial and can take a large percentage of the time dedicated to the preference elicitation procedure, making them potentially useless in a very large number of practical situations, stakeholders' workshops included (Meyer and Ponthière 2011).

Index[4]), such an additive representation is likely to be inadequate for the purposes at hand, because this requires individual preferences over multiattribute societies to satisfy the assumption of *mutual preferential independence* among all the attributes (see below for a definition). Since the former is a very strong requirement which is likely to be violated by individual preferences on multiattribute societies, it makes sense to allow a priori the possibility of interactions between the various dimensions or attributes of societies. A natural way to take into account both the importance of each attribute and subset of them is to consider the representation of individual preferences by means of the Choquet integral aggregator. This is the natural extension of the weighted arithmetic mean where integration is defined with respect to a non-additive measure rather than an additive one. In this section we only describe the basic features of the aforementioned methodology. A more technical treatment can be found in the "Appendix 1".

The structure of MAVT with choquet integration

Let $X \subseteq X_1 \times X_2 \times ... \times X_n$ with $n \geq 2$ be a set of objects described by a set $N := \{1, ..., n\}$ of decision attributes. For example, it could be a set of Science Parks each characterized by n performance attributes. We consider now a fictitious decision maker (DM), whose preferences, expressed by a binary relation $\succeq$ on X, can be represented through a value function $U : X \rightarrow \mathbb{R}$ such that

$$x \succeq y \iff U(x) \geq U(y) \quad \forall x, y \in X.$$

In our case, the decision objects, i.e. the elements in the set X, consist of Science Parks, whereas the attributes under study, i.e. components of each vector $x \in X$, represent different characteristics or dimensions of performance associated with each SP (e.g. sales' growth of incubated firms or number of research centres hosted by the park). The value function U is retrieved through an interactive and incremental process requiring the DM to express his/her preferences over a small subset of selected objects. Hence, it is possible to consider the function U as a numerical representation of the preference relation $\succeq$ defined on X and to use it as a *model* for DM's preferences.

In this study we consider the general *transitive decomposable model* of Krantz et al. (1971) where it is possible to define

$$U(x) := F(u_1(x_1), ..., u_n(x_n)) \quad \forall x = (x_1, ..., x_n) \in X$$

where the functions $u_i : X_i \rightarrow \mathbb{R}$ are called *marginal value functions* and $F : \mathbb{R}^n \rightarrow \mathbb{R}$, non decreasing in all its arguments, is called the *aggregation function*. Under this framework it would be possible to interpret each quantity $u(x_i)$ as a measure of the satisfaction of the DM along attribute (or dimension) $i \in N$. The exact form of U is case-dependent.

A particular case of value function arises under the assumption of *mutual preferential independence*:[5] when it can be assumed that the value function becomes additive so to

---

[4] The PPI (Fish 2006), an index used to assess the functions and independent position of the parliament, as well as the Severity Index (Reinhart and Rogoff 2014), which refers instead to the magnitude of recession episodes, represent additional examples of naïve methods of aggregation that are relatively common in the economics and politics literatures.

[5] Attributes x_1 and x_2 are mutual preferential independent if $(x_1, x_2^0) \succ (x_1', x_2^0)$ then $(x_1, x_2) \succ (x_1', x_2) \, \forall x_2$ and $(x_1^0, x_2) \succ (x_1^0, x_2')$ implies $(x_1, x_2) \succ (x_1, x_2') \, \forall x_1$.

obtain $U(x) = \sum_{i=1}^{N} u_i(x_i)$ (Debreu 1960). It is easy to see that this representation coincides exactly with a weighted average.

However, when interaction phenomena among attributes do not want to be excluded a priori, it has been proposed to substitute the weight vector involved in the calculation of weighted sums with a monotone set function on N, called *capacity* (Choquet 1953) or fuzzy measure (Sugeno 1974), which might be non-additive. This allows to take into account not only the importance of each attribute for the DM but, also, the importance of each subset of attributes and possible complementarity or redundancy among them. In such a context, a natural extension of the weighted arithmetic mean is the Choquet integral with respect to a capacity. To be more precise, a *capacity* is simply defined as a set function $\mu : \mathcal{P}(N) \rightarrow [0, 1]$ such that $\mu() = 0$ and $\mu(N) = 1$. Moreover, for any two subsets $A, B \subseteq N$ such that $A \subseteq B$, we get $\mu(A) \leq \mu(B)$. When a capacity satisfies some additional conditions, we can say that it is *additive* and, in these cases, it corresponds to a probability measure. However, in general, this is not the case and, as an immediate consequence, it becomes important to recall that the importance attributed to joint performances along different dimensions is different from the sum of the importance attributed to the same dimensions individually.

The notion of capacity μ on N is crucial within this setting and leads to the definition of the so-called *Choquet integral* in the context of MAVT.

Definition   The Choquet integral of an alternative x, represented by the vector of partial values, $u(x) := (u_1(x_1), ..., u_n(x_n))$ w.r.t. a capacity μ on N is defined by

$$C_\mu(u(x)) := \sum_{i=1}^{n} u_{\sigma(i)}(x_{\sigma(i)})[\mu(A_{\sigma(i)}) - \mu(A_{\sigma(i+1)})]$$

where σ is a permutation on N such that $u_{\sigma(1)}(x_{\sigma(1)}) \leq ... \leq u_{\sigma(n)}(x_{\sigma(n)})$. Also, $A_{\sigma(i)} := \sigma(i), ..., \sigma(n), \forall i \in 1, ..., n$, and $A_{\sigma(i+1)} := \emptyset$.

The Choquet integral acts as an aggregation operator with respect to μ that accounts for the role played by each subset of attributes in the decision problem. The standard weighted arithmetic mean coincides with the Choquet integral in the special case of an additive capacity, which implies the independence of the attributes. From a behavioural point of view the use of weighted averages is equivalent to assume that the DM only considers attributes on their own and not in groups. As we have stressed so far, this might well not be the case.

In order to provide a more manageable, representation of the Choquet integral, we introduce the so-called *Möbius transform* and the notion of *k-additivity*. While the first is only a different representation of the capacity μ, the concept of *k-additivity* is rather crucial as it captures the important trade-off between the complexity of the capacity and its modelling representability. In particular, it indicates the minimum cardinality of the sets of attributes that need to be used to represent a specific preference relation $\succeq$ (which takes the form of a partial weak order). For example, a *k-additivity* of 3 indicates that the preferences of a DM cannot be represented by a model like the one introduced so far if the interactions among sets of 2 and 3 attributes are not considered. Obviously, a *k-additivity* of 1 corresponds to an underlying model that takes the form of a weighted sum, as only single attributes can be used to represent preferences. For a formal definition of the Möbius transform of a capacity and the notion of *k-additivity* refer to the "Appendix 1".

Analysis of the aggregation

As previously mentioned, the use of the Choquet integral as an aggregation operator allows to account for the interaction among attributes. This is due to the fact that a weight of importance is attributed to every subset of criteria rather than to each criterion taken on its own. However, it is not a trivial exercise to resort these features from definition. To overcome this difficulty different indices has been proposed in the literature (Grabisch 1996; Marichal 2000) and two of them are particularly useful in our context. Their formal definitions, expressed by means of the capacity μ, can be found in the "Appendix 1".

Importance index

This index is used to denote the overall importance assigned to a single attribute (that will be, in our case, an indicator of SPs' performances) in a decision problem where alternatives are characterized multiple attributes. In the context of cooperative game theory, Shapley (1953) introduced a particular coefficient of importance that exactly serves to the scope. It is obtained by averaging all the marginal gains obtained by adding the criterion to every group (or coalition) not including itself. As shown in the "Appendix 1" the marginal gain for each attribute i can be expressed in terms of the capacity μ and corresponds, roughly speaking, to the difference between the importance that the DM assigns to a coalition of attributes where i is included and that assigned to the group formed by the same attributes but i.

Interaction index

While it is very useful to characterize the importance of each attribute in a multidimensional decision problem, the Shapley values tells nothing about the value provided by jointly scoring well (or bad) along groups of different attributes. Putting it differently, it does not provide any information on the interaction effects. Indeed, consider for instance two attributes i and j such that $\mu(ij) > \mu(i) + \mu(j)$ and recall that $\mu(\cdot)$ can be interpreted as the relative importance assigned to a coalition of attributes. This clearly shows a *complementarity* effect between the two, that is, the value they provide together is larger than the sum of the values they are able to provide individually. Similarly, the inequality $\mu(ij) < \mu(i) + \mu(j)$ models a *redundancy*. Finally, if the two attributes i and j do not interact at all we clearly have $\mu(ij) = \mu(i) + \mu(j)$, which means that i and j play independent roles. Therefore, assuming that i and j are positively correlated or complementary (resp. negatively correlated or redundant), then the marginal contribution of j to every combination of attributes that contains i should be strictly greater than (resp. less than) the marginal contribution of j to the same combination when i is excluded. To quantify the overall degree of interaction between any pair of attributes we use an *Interaction index* originally proposed by Murofushi and Soneda (1993) and obtained by the average value of these marginal contributions.

The capacity identification

As the Choquet integral involves a capacity that is defined by $2^n - 2$ coefficients,[6] it is difficult that the decision maker is able to provide these parameters directly and therefore some data are necessary to infer the underlying capacity. To this purpose, we firstly determine the marginal value function of each decision maker; secondly, we identify a

---

[6] Corresponding to the total number of possible groups of attributes less the empty and full sets.

capacity, if it exists, such that the Choquet integral w.r.t. this capacity numerically represents the preferences of the decision maker.

The preferential information expressed by the decision maker are supposed to rely on a finite and usually small subset O of the set X of objects of interest, where the set O is usually composed either of available objects or of selected, potentially fictitious objects. In our case the objects of interest are represented by fictitious Science Parks, characterized by different multidimensional performances.

As described in Grabisch et al. (2008), once an appropriate subset O has been determined, each subject involved in the decision problem is asked to express her *initial preferences*. The initial preferences, from which the capacity will be determined, can take the form of:

- a partial weak order $\succeq_O$ over O (ranking of the available objects);
- a partial weak order $\succeq_N$ over N (ranking of the importance of the attributes);
- a partial weak order $\succeq_P$ over P on the set of pairs of attributes (ranking of interactions);
- etc.

In line with the existing literature, the identification method can be expressed as an optimization problem, where the initial preferences of the decision maker define the constraints. There are various optimization approaches that differ according to their objective function and the preferential information they require as input. In our setting, we rely on a method based on *minimum variance* identification principle, whose main idea is to favour the "least specific" capacity, if any, compatible with the initial preferences of the decision maker. The most relevant advantage of this approach is that it leads to a unique solution, if any, because of the strict convexity of the objective function. Moreover, in the case of initial preferences that involve a small number of constraints ("poor"), this unique solution will not exhibit too specific behaviours characterized for instance by very high positive or negative interaction indices or very uneven Shapley values (Grabisch et al. 2008). In addition, the minimum variance approach in our case allows the best comparison with the simple average as it minimizes the distance from a uniform distribution. The use of a uniform as a benchmark appears a reasonable choice in a context where, *a priori*, there is no reason to assign different importance to different attributes of performance[7] and, if a difference is entailed in a subject's preferences, we use the least polarizing representation compatible with the observed difference. Notwithstanding this conservative choice, in "Results I: behavioural analysis of the aggregation" section we show that subjects actually assign remarkably different importance to our performance dimensions.

It is relevant to recall that a solution to our identification problem is a general capacity defined by $2^n - 1$ coefficients, which completely characterizes each DM. However, the same preferences can be represented through different k-additive capacities, where $k \in 1, .., n$ and, typically, $k = 2$ or 3. In our setting, we choose the numerical representations with the lowest possible level of k-additivity. This choice is justified in two ways. First, we check whether a suitable additive model (1-additivity) exists and, if preferences are so complex and interactions so relevant that it cannot be the case, then we select the simplest possible representation.[8]

---

[7] In a sense, we invoke the principle of insufficient reason Bernoulli (1713).

[8] Notice that this approach is quite in line with the literature. Look for example at Meyer and Ponthière (2011) and references therein.

A pilot study: application to Italian Science Parks

In this section we present the setting of our study, which aims at evaluating and comparing performances of 56 Italian Science Parks. As previously mentioned, SPs' stakeholders are variegate and comprehend both single individuals (e.g. entrepreneurs or parks' managers) and more complex entities or institutions (e.g. local administration or universities). A comprehensive evaluation and comparison of SPs performances would need to properly represent each stakeholder category in the decision problem and, possibly, to involve them in the choice of performance attributes. Being ours a pilot study and a completely novel approach in the field of SPs' evaluation, we simplify the setting by considering only two classes of stakeholders and relative small set of attributes. Despite this, the procedure we apply can be straightforwardly repeated with a much larger number of stakeholders and dimensions characterizing SPs.

The starting point of the analysis consists of submitting to two groups of respondents a standardised questionnaire asking them to rank hypothetical multidimensional Science Parks. In particular, we chose two categories of SPs stakeholders: (1) students of entrepreneurship and innovation management as potential entrepreneurs and (2) university researchers. [9] The first group refers to 20 students from two MSc programs in Entrepreneurship (10) and Innovation Management (10).[10] The second group, instead, is composed by academic researchers from Bocconi University, all having at least 1 year of research experience within the field of technology transfer, firm innovativeness or regional innovation systems.[11] Each respondent is iteratively treated as the decision maker described above and, on the basis of her answers, we elicit individual preferences via a Choquet integral-based MAVT model.

The discussion concerning the attributes of performance we use to evaluate and compare each Italian SP is devoted to the next subsection.

Dimensions and attributes

Considering that there is a large number of dimensions characterizing SPs' activities, selecting the most relevant or appropriated is not a trivial task. The existing literature suggests a plethora of different indicators, reflecting the degree of fulfilment of various objectives attributable to SPs. Past evaluations mainly assess economic performance of tenant firms' using the following indicators as measures of a successful development program: employment, value added, survival rate and the number of jobs created (Luger and Goldstein 1991). Monck (2010) divides performance indicators into three sub groups (key performance indicators, intermediate results and short term indicators) and, remarkably, adds features of the SP itself to the evaluation setting. For example, he proposes to consider the number of available slots to host firms and the number of connections created with knowledge based collaborators. Dabrowska (2011) presents a nice overview of the

---

[9] Note that in different studies both actual/potential entrepreneurs, students and university research personnel have been identified as relevant SPs stakeholders (Vedovello 1997; ANGLE 2003; Link and Scott 2003; Hansson et al. 2005).

[10] The first program has been held at Bocconi University (Milan) while the second at Insubria University (Varese); in both cases students have attended a course on technology transfer and university-industry relations before answering the questionnaire.

[11] In particular the group is composed by 1 full professor, 1 lecturer, 1 assistant professor, 1 Ph.D. candidate and 6 graduate research assistants.

SPs' evaluation literature and collects results of a workshop held by the IASP[12] on performance measuring. In particular, she details a long list of dimensions, encompassing almost each activity a SP might entertain. What appears to be in common is the idea that SPs have to be evaluated accounting for their double nature, considering their role as entrepreneurship supporting organizations (according to the IASP website "they should facilitate the creation of new businesses via incubation and spin-off mechanisms, and accelerate the growth of small and medium size companies") and as innovation inducing institutions ("stimulate and manage the flow of knowledge and technology between universities and companies and provide environments that enhance a culture of innovation, creativity and quality"). On the other side, the possibility of appropriately comparing different SPs is constrained by data availability problems (Hodgson 1996, see also). Publicly available datasets seem not to exist (to the best of authors' knowledge) and the majority of studies relies on survey-extracted information.[13] Notwithstanding this limitation we have selected a pool of eight performance indicators that will be used to provide a pilot comparison of all the 56 Italian SPs we have information about (Fig. 1). Remarkably, all these attributes (but one, entropy) belong to the list identified in Dabrowska (2011), with some of them included also in Monck (2010). In addition, we have kept separated the sets of indicators referring to the *entrepreneurrhip* and *innovation* SPs' natures outlined above, with 4 attributes characterizing each of the two. The resulting tree of indicators allow us to investigate redundancies and complementarities among dimensions of performance both within and between the two branches. Specifically, the innovation dimension comprehends the number of research centres hosted by the park, an indicator of firms' patenting activity (cumulative number of patent firms applied to in the period 2010–2012[14]), the number of links with universities and the number of research projects the park is involved in. On the other hand attributes charactering the entrepreneurship dimension are: the rate of growth of affiliated firms in the period 2010–2012 (measured as the logarithmic difference of gross sales between two consecutive periods), the number of job-places created after firms' establishment (employees), the average distance between firms affiliated to the park but not located therein and the park itself (which could be thought as a proxy for knowledge spillovers from affiliated firms), and finally the degree of specialization of the park (which is measured trough an entropy coefficient of the distribution of firms along the industries they belong to[15]).

Fictitious science parks

The first stage of our evaluation procedure consist in the elicitation of preferences from the set of stakeholders. This step is carried out in a very intuitive way, by firstly asking respondent to rank different science parks according to their own preferences and then

---

[12] International Association of Science Parks, http://iasp.ws.

[13] Also the data about SPs and their tenants used in this paper largely come from a survey conducted in 2012 and submitted to all Italian SPs. Survey data have then been matched with corporate data from the Bureau van Dijk's ORBIS database and the PATSTAT database (Ferrara et al. 2012; Lamperti et al. 2015, for additional information).

[14] This time span has been chosen since the last SPs in our sample has been established in 2010 and we did not want to exclude it from the evaluation; moreover, we have planned to enrich and repeat the evaluation every three years to monitor how the set of Italian SPs perform in a reasonable time horizon for parks' activities.

[15] Industries are identified through the NACE Rev. 2 classification.
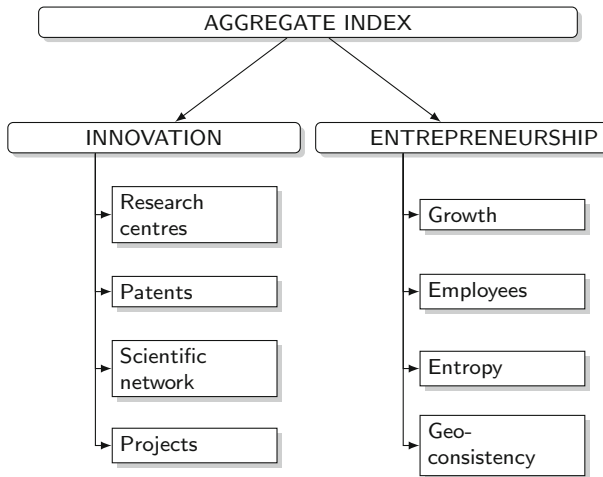
Fig. 1 The tree of performance indicators

solving the corresponding capacity identification problem outlined in "The capacity identication".

With reference construction of multidimensional Science Parks to be ranked, it is important to notice respondents were asked to rank hypothetical societies, whose performances on each dimension under study did not necessarily coincide with any observed real structure. This choice guarantees the advantage of preventing subjects from favouring specific SPs on the basis of features which are not explicitly considered in our approach (e.g. geographical location). However, an important step consists in restricting the set of decision objects to a subset of relatively plausible SPs that are not "too different" from the existing ones.

Accordingly, hypothetical SPs are here constructed by departing from a *SP of reference*, whose attributes take levels that are in line with the prevailing performance levels registered in 2012 by Ferrara and Mavilia (2012). In particular, such levels are normalized on a [0, 100] scale where 100 is assigned to the SP displaying the highest score on the particular dimension considered.[16] The reference SP is then assigned on each of the eight dimensions the average score registered among Italian SPs. Besides the reference point, four additional levels of achievements are then introduced. The level "good" amounts to an achievement of 110 % of the standard achievement, similarly "very good" amounts to 120 %, "bad" to 90 % and "very bad" to 80 %. The choice of an interval defined by a *plus* or *minus* 20 % around the reference point depends on the need to include SPs that are plausibly reachable (Meyer and Ponthière 2011). Table 1 presents the outcome of the construction of hypothetical SPs, considering the two main dimensions and the levels of the related attributes.

Respondents were asked to provide a series of rankings on different sets composed of five fictitious SPs.[17] Each set is characterized as follows: the SPs to be evaluated vary

---

[16] The normalization procedure is simple and allows to remove issues due to units of measurement. Therefore, the relative scores of each SPs on different dimensions of performance (e.g. number of job created and sales growth) are directly comparable.

[17] Specifically, they were asked to provide an order, assigning 1 to the SP which is though of be the best performing and 5 to the worst.

Table 1 Attributes' levels of hypothetical science parks

Attributes	Very bad SP (−20 %)	Bad SP (−10 %)	SP of reference	Good SP (+10 %)	Very good SP (+20 %)
Innovation					
Research centres	12,5	14,1	15,7	17,2	18,8
Patents	6,9	7,7	8,6	9,4	10,3
University	17,9	20,1	22,3	24,6	26,8
Projects	11,7	13,2	14,6	16,1	17,5
Entrepreneurship					
Growth	12,4	13,9	15,5	17	18,6
Employees	10,4	11,7	13	14,3	15,6
Entropy	45,2	50,8	56,5	62,1	67,8
Distance	23	25,9	28,8	31,7	34,5

along two of the possible attributes, while all the others are kept fixed to allow the subject concentrating on the relative importance of pairs of attributes taken in isolation. After this preliminary phase, the best SPs identified in each set have been grouped together and the respondent asked to provide a new ranking on all these fictitious SPs, whose performances are now likely to vary along each attribute. This procedure has been carried out for each respondent at all possible nodes of the decision three, which are, in our case, the entrepreneurship area, the innovation area and the aggregate index. Details about the procedure and the questionnaire are included in the supplementary material.

Marginal value functions

A key decision in the setting of the capacity identification problem consists in the choice of the marginal value function for each respondent (see "Methodology" section). The shape of the marginal value function reflects the behavioural attitudes of the respondent, for example it determines whether she values losses and gains symmetrically from the status quo or gives more importance to one of the two. The main novelty of our analysis consists in directly assessing its shape rather than arbitrarily assuming the standard symmetric S-shape form (Meyer and Ponthière 2011, used for example in). Allowing for this flexibility is particularly relevant when different stakeholders (e.g. park managers, entrepreneurs, researchers) are involved in the same decision problem. To recover marginal value functions in a practical and consistent way with what can be asked to respondents in reasonable time, we apply the methodology originally developed by Kahneman and Tversky (1979). In their *prospect theory*, they find an asymmetric s-shaped marginal value function to be more realistic, and value is assigned to gains and losses rather than to final assets. Specifically, value should be treated as a function in two arguments: the asset position that serves as reference point, and the magnitude of the change (positive or negative) from that reference point. They assume that the individual response is a concave function of the magnitude of monetary changes. Thus, the difference in value between a gain of 100 and a gain of 200 appears to be greater than the difference between a gain of 1100 and a gain of 1200. The same reasoning holds for losses rather than gains. Therefore, they hypothesize that the value function is concave above the reference point ($u''(x) < 0$, for $x > 0$) and convex below it ($u''(x) > 0$, for $x < 0$). This means that the marginal value of both gains and losses generally decreases with their magnitude. Strictly

speaking, the value function is defined on deviations from the reference point; it is normally concave for gains, commonly convex for losses; and it is generally steeper for losses than gains. This last characteristic is determined by the fact that the aggravation experienced during a loss appears to be larger than the satisfaction associated with a gain equal in magnitude. In order to choose a suitable function for each respondent, we design a lottery similar to that of Kahneman and Tversky (1979) with three questions related to possible interventions which may modify a SP's overall performance. We then use the link between the shape of value functions and preferences over lotteries as a driver for the choice of a suitable function.

The detailed description of the lottery questions are specified in the supplementary material. However, we recall here that respondents have been asked to put themselves in the shoes of a SP's manager and think about which action to take in three different problems entailing the possibility to improve the performance of the park or, in case of failure, to dampen it. The outcomes of each problem together with the probabilities of success are reported here (the first number indicate the outcome with the second its occurrence probability):

1. L1: (30, 0.25) versus L2: (20, 0.25; 10, 0.25)
2. R1: (−30, 0.25) versus R2: (−20, 0.25; −10, 0.25)
3. S1: (10, 0.50; −10, 0.50) versus S2: (30, 0.50; −30, 0.50).

Respondents' answers allowed us to deduce the shape of the possible marginal value function representing subjects' preferences for each attribute of performance. In particular, question 1. and 2. allowed to deduce the shape of the function in the positive and negative domain respectively, while question 3. has been used to infer the relative steepness of the two sides (Kahneman and Tversky 1979, for details). Accordingly, Table 2 defines the possible different shapes of the candidate functions and presents the number of respondents associated to each of them, showing that the majority is characterized by the s-shape loss averse but the degree of heterogeneity is quite large and cannot be overlooked.

Table 2 Value functions association

Question 1	Question 2	Question 3	Shape	Associate function	# Respondents
L1	R1	Any	Convex	Pessimistic	3
L2	R2	Any	Concave	Optimistic	4
L1	R2	Any	Reverse s-shaped	Reverse s-shaped	2
L2	R1	S1	S-shaped	S-shaped loss averse	11
L2	R1	Indiff	S-shaped	S-shaped	3
L2	R1	S2	S-shaped	S-shaped gain lover	6
Indiff	Indiff	any	Linear	Linear	0
Indiff	R1	Any	Positive semilinear s-shaped	Loss pessimistic	0
Indiff	R2	Any	Positive semilinear reverse s-shaped	Loss optimistic	0
L1	Indiff	Any	Negative semilinear reverse s-shaped	Gain pessimistic	0
L2	Indiff	Any	Negative semilinear s-shaped	Gain optimistic	1
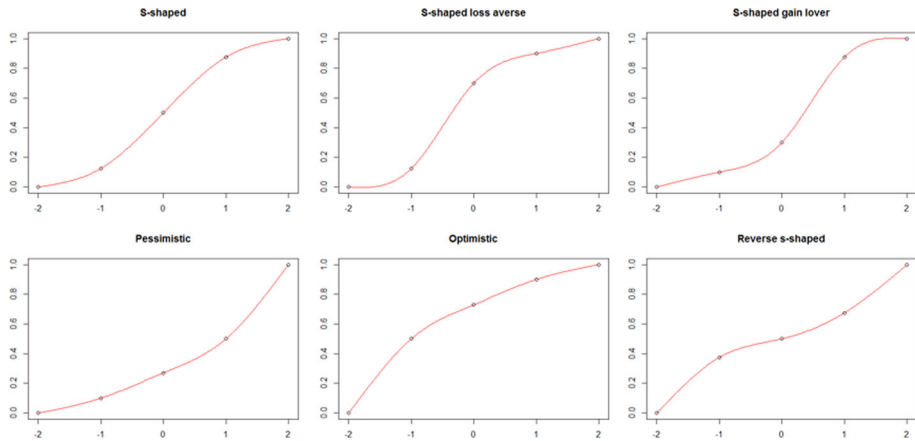
Fig. 2 Marginal value functions

Moreover, Fig. 2 shows the shapes of the most frequent value functions, where the y axes refers to the value associated to a performance quantified in x and 0 stands for the reference point. It is clear to see how the postulation of a symmetric s-shaped for all subjects , such as in Meyer and Ponthière (2011), cannot be a good approximation to our context, where 27 over 30 respondents' answers are not consistent with such a shape.

Results I: behavioural analysis of the aggregation

The procedure we have described so far provides two distinct kind of results. On one side, it aggregates preferences and directly computes an index, the Choquet integral, that can be used to compare each structure in our sample and look for best and worst performers. On the other side, it allows to analyse the aggregation from a behavioural point of view, shading lights on how stakeholders (individually and on aggregate) implicitly assign importance and (possibly) interactions to the different attributes of performance. This section is devoted to the discussion of such features, while the next one treats the comparison of actual SPs.

After having collected information about stakeholders' preferences and elicited them through the methodology outlined in "Methodology" section, the most interesting part of the study steps in and consists in the analysis of the aggregation. Note that, for practical purposes, capacity elicitation problems can be solved using and adapting the routines implemented in the Kappalab R package (Grabisch et al. 2008).[18]

First, we provide some evidence that using an additive model to compare complex institutions like Science Parks is likely to be inadequate. Table 3 shows the number of respondents associated to each level of k-additivity. It immediately emerges that out of 90 potential outcomes, $k \neq 1$ that corresponds to a non-additive model, holds for more than

---

[18] This package allows the determination of a set of marginal value functions which are compatible with the respondents' ranking, to test the presence of an additive numerical model and to calculate the corresponding capacities.

Table 3 K-additivity

k-additivity	# Respondents		
	Innovation	Entrepreneurship	Final
1	10	14	16
2	13	9	12
3	7	7	–
4	0	0	–

half of the subjects.[19] Specifically, results show that only 2 out of 30 respondents display $k = 1$ in all three nodes of analysis, while 8 out of 30 in 2 out of 3 nodes. This result fully justifies our approach that considers a linear additive model to be inappropriate for representing stakeholders' preferences. Moreover, it is interesting to notice that some differences already emerge between the two categories of stakeholders we consider: the additivity model mainly holds for students, while higher levels of k-additivity are common among researchers, who seem to hold more complex preferences about what renders a SP better than another.

Differences among stakeholders: students versus researchers

When different stakeholders are looking for a best performer, each might take into different consideration the various activities entertained by SPs. For example, a policy-makers might look at creating job opportunities within incubated firms differently from a researcher would do. Our approach offers a flexible environment to investigate the presence and analyse such differences.

In the context of our pilot study, let us focus on the relative importance assigned to each performance attribute by the two stakeholder categories. For that purpose, Table 4 presents the values of the Shapley indexes for each students, while Table 5 the researchers' ones. On the one hand, researchers are both involved in the evaluation of SPs and, potentially, might find job-opportunities within the park; on the other hand, students can be regarded as potential entrepreneurs. It is therefore interesting to see how they differently consider the importance of SPs' attributes. Recall that the Shapley value of an attribute stands for the average value of the marginal contribution of that attribute to a subset of attributes not containing it.

Both tables invite two main observations. First, the indexes take generally very different values, suggesting that indicators of performance assigning equal weights to all dimensions of SP misrepresent the inherent complexity of multi-attribute structures. Second, a significant heterogeneity emerges across respondents. Looking at the averages in the bottom row of both tables, it turns out that for the innovation area research centres and projects are more important for researchers while patenting activity and scientific networks are more relevant for students. This result is not surprising. Being part of the everyday life of the university, students underline the importance of the links with the academia as well as the significance of innovation outcomes, i.e. patents. Conversely, researchers are more interested in the research environment behind the innovation outcome, that is the existence of research centres and projects aiming at developing new processes and products. However,

---

[19] Note that the last column does not sum to 30 because for 2 respondents the identification of the final capacity was not possible with any level of k-additivity. Therefore, they have not been considered in the final aggregation (see "Properties of the aggregation" section).

Table 4 The values of the Shapley indexes for the 20 students

Resp	Innovation				Entrepreneurship			
	Research centers	Patents	Scientific network	Projects	Growth	Employees	Entropy	Geo-consistency
1	0.26	0.13	0.42	0.19	0.04	0.02	0.27	0.13
2	0.26	0.18	0.38	0.18	0.34	0.25	0.22	0.19
3	0.15	0.55	0.02	0.28	0.55	0.15	0.28	0.02
4	0.13	0.53	0.30	0.03	0.55	0.15	0.15	0.15
5	0.26	0.18	0.38	0.18	0.29	0.12	0.29	0.29
6	0.27	0.11	0.43	0.19	0.16	0.14	0.42	0.28
7	0.25	0.32	0.25	0.19	0.55	0.02	0.28	0.15
8	0.27	0.13	0.00	0.00	0.02	0.13	0.04	0.27
9	0.00	0.00	0.13	0.27	0.08	0.26	0.36	0.30
10	0.22	0.55	0.02	0.22	0.55	0.15	0.28	0.02
11	0.22	0.55	0.02	0.22	0.41	0.05	0.27	0.27
12	0.32	0.10	0.39	0.20	0.55	0.28	0.15	0.02
13	0.30	0.30	0.12	0.28	0.43	0.20	0.19	0.18
14	0.29	0.18	0.18	0.35	0.17	0.27	0.03	0.27
15	0.47	0.31	0.15	0.07	0.43	0.27	0.19	0.11
16	0.00	0.23	0.17	0.00	0.04	0.02	0.27	0.13
17	0.22	0.55	0.02	0.22	0.24	0.37	0.16	0.24
18	0.28	0.17	0.09	0.46	0.33	0.25	0.37	0.05
19	0.38	0.15	0.26	0.21	0.38	0.24	0.21	0.18
20	0.22	0.31	0.22	0.24	0.25	0.16	0.32	0.25
Average	0.24	0.28	0.20	0.20	0.32	0.18	0.24	0.17

Table 5 The values of the Shapley indexes for the 10 researchers

Resp	Innovation				Entrepreneurship			
	Research centers	Patents	Scientific network	Projects	Growth	Employees	Entropy	Geo-consistency
21	0.08	0.21	0.21	0.50	0.45	0.05	0.25	0.25
22	0.55	0.15	0.02	0.28	0.55	0.15	0.28	0.02
23	0.16	0.33	0.05	0.45	0.49	0.30	0.08	0.13
24	0.13	0.49	0.08	0.29	0.49	0.30	0.13	0.08
25	0.11	0.27	0.19	0.43	0.43	0.27	0.19	0.11
26	0.39	0.06	0.39	0.17	0.49	0.10	0.30	0.10
27	0.35	0.22	0.28	0.13	0.15	0.02	0.55	0.28
28	0.45	0.05	0.33	0.16	0.25	0.45	0.05	0.25
29	0.55	0.22	0.22	0.02	0.55	0.22	0.22	0.02
30	0.31	0.31	0.07	0.31	0.39	0.22	0.23	0.16
Average	0.31	0.23	0.18	0.27	0.42	0.21	0.23	0.14

note that a high variability is present within groups. For instance, students 1 and 2 assign a large weight to research centres and a low weight to patents, whereas the opposite occurs for respondents 3 and 4.

As for the entrepreneurship area, both categories assign on average the highest importance to growth, however the magnitude of the average index is larger for researchers. Similarly, both researchers and students share the lowest values for the geo-consistency criterion. The number of employees and the degree of entropy lie in between, displaying very similar indexes across both categories.

In order to further investigate this heterogeneity, we suggest to perform a cluster analysis with respect to the elicited capacities that allows to empirically test whether specific patterns of preferences can be identified within different stakeholders' categories. In order to do that, we rely on hierarchical clustering (also called hierarchical cluster analysis or HCA), that is a method of cluster analysis seeking a hierarchy of clusters. Specifically, we adopt a *agglomerative* strategy, which corresponds to a "bottom up" approach where each observation starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy.

In order to decide which clusters should be combined, a measure of dissimilarity between sets of observations is required. In most methods of hierarchical clustering, this is achieved by use of an appropriate metric (a measure of distance between pairs of obser-vations), and a linkage criterion which specifies the dissimilarity of sets as a function of the pairwise distances of observations in the sets. Here we adopt the Euclidean distance as metric, which is commonly defined as

$$\| a - b \|_2 = \sqrt{\sum_i (a_i - b_i)^2}$$

and the *Ward's criterion* as linkage criterion that includes the decrease in variance for the cluster being merged. Ward's minimum variance criterion minimizes the total within-cluster variance where, at each step, the pair of clusters with minimum between-cluster distance are merged. To implement this method, at each step we need to find the pair of clusters that leads to the minimum increase in total within-cluster variance after merging. This increase is a weighted squared distance between cluster centers. At the initial step, all clusters are singletons (clusters containing a single point).

Figure 3 presents three hierarchical dendograms[20] that graphically illustrate the levels of subjects' aggregation at each node of the attributes' tree (1) that are labelled *innovation*, *entrepreneurship* and *final*. Recall that subjects labelled by a number between 1 and 20 are students, while those in the range 21–30 are researchers.

The appropriate number of clusters was chosen implementing the *k-means algorithm* so that the trees have been cut in five main clusters (red rectangles). The overall picture suggests that for innovation (a) and entrepreneurship (b) some clustering effect between students and researchers exist. For instance, in (a) we have only students in the second and fourth cluster (counting from left to right), while only researchers in the last one. Con-versely, for the entrepreneurship area researchers are concentrated in the first three clusters. As for the final index (c), clusters' composition is quite mixed and there is none totally composed either by students or researchers. These results show that for the two lower nodes, entrepreneurship and innovation, the "type" of stakeholder plays a role; there are different small groups of subjects sharing similar opinions which are composed solely by

---

[20] The dendogram is a tree diagram frequently used to illustrate the arrangement of the clusters produced by hierarchical clustering.

Fig. 3 Cluster analysis:
hierarchical dendrograms.
a Innovation. **b** Entrepreneurship.
c Final

(a) Innovation

(b) Entrepreneurship

(c) Final
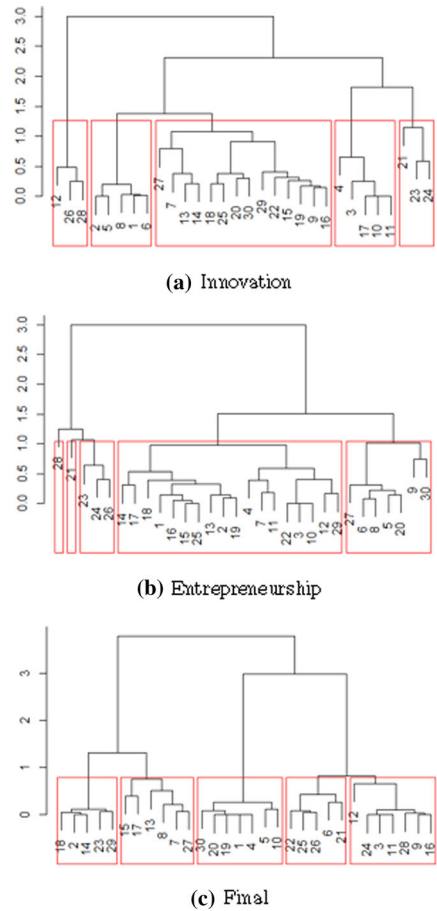
students or researcher and a large mixed group, meaning that similar respondents who are extremal with respect to the average tend to belong to the same group. While the majority of respondent assign similar importance and interaction to the same set of attributes there are small coalitions of either students or researchers each focusing on the importance of different sets of attributes.

Properties of the aggregation

Once individual preferences have been elicited they should also be aggregated in a single capacity, which is then used as a summary measure to determine the aggregate importance and interactions of different performance attributes had all respondents been considered as a unique body. The aggregation of preferences is everything but a straightforward task, which has already been addressed in various fields, going from public economics (Smith 1973) to engineering (Moon and Kang 1999), game (Gerardi et al. 2009) and decision theory (Hsu and Chen 1996). In the context of multi-criteria decision making, where our approach falls within, a popular solution to the problem consists in weighted aggregation, where weights depends upon features of the evaluating subjects. In a recent application

Pinar et al. (2014) assigned weights ex-post, after having observed how close different respondents are in a suitable multidimensional space. Our approach takes an opposite view and assigns equal weights ex ante, that is, before observing how close respondents are. This is motivated by the fact that, in our context, we do not see any sufficient reason to justify the choice of more important respondents relatively to the others. Formally, each stakeholder i is fully characterized at the node $j = \{\text{INN}, \text{ENTR}, \text{FINAL}\}$ of the attributes' tree by the fuzzy set

$$A_{i,j} = \{(x_j, \mu_{i,j}), x_j \in X_j \equiv 2^{N_j}\} \tag{1}$$

where N_j is the set of attributes under node j and $\mu_{i,j} : X_j \to [0, 1]$ is the elicited capacity function. Then, the aggregate capacity can be written as

$$\overline{\mu}_j = \sum_i \lambda_i \overline{\mu}_{i,j} \tag{2}$$

where $\overline{\mu}_{i,j}$ is the vector collecting values elicited from subject i for each attributes' set available at j and λ_i is the weight associated with respondent i. In our particular case, where all stakeholders are given the same importance, $\lambda_i = 1/M \quad \forall i$, where M is the total number of subjects involved.

Starting from the aggregate capacity $\overline{\mu}_j$ as input, a number of results about the behavioural features of the evaluation procedure can be extracted and discussed. Let us begin with the Importance indexes. Table 6 presents both the relative (branch-specific) and the global Shapley values of the aggregated decision maker. The global importance index can be obtained by multiplying the Shapley values along the branches of the aggregation tree. These results are also directly compared to the weights assigned to the same attributes by Ferrara and Mavilia (2012), who used the same SPs data as we do but applied a linear aggregation with arbitrarily chosen weights.

According to the specific set of stakeholders involved in this study, it emerges a prevalence of the innovation dimension over the entrepreneurship one in evaluating SPs performances. This feature is visible from the comparison of the respective importance indexes (0.62 v.s. 0.38). Such a result falls in line with the academic literature, which tend to primarily evaluate SPs on the basis of their role as innovation inducing and research stimulating organizations (Felsenstein 1994; Colombo and Delmastro 2002; Lamperti et al. 2015; Minguillo et al. 2015). Among performance attributes included in the innovation branch, research centres and patents turn out to be the most relevant, although the

Table 6 Shapley values for the aggregated game	Relative	Global	FM (2012)
Innovation 0.62			
Research centers	0.29	0.18	0.20
Patents	0.27	0.17	0.20
Scientific network	0.21	0.13	0.10
Projects	0.24	0.15	0.20
Entrepreneurship 0.38			
Growth	0.38	0.14	0.20
Employees	0.20	0.08	0.03
Entropy	0.25	0.09	0.03
Geo-consistency	0.16	0.06	0.03

magnitude of the four coefficients is quite similar. This result suggests that the higher importance of innovation is not driven by a single attribute but it is rather a shared effect. Moving to the entrepreneurship area, we find a strongly leading criterion (firms' growth of sales), with a Shapley value of 0.38, whereas the lowest value is taken by geo-consistency (relative importance index of 0.16). What turns out to be surprising concerns the importance that our stakeholders assigned to jobs' creation as a measure to assess SPs' success. While the early literature on SPs' performance evaluation attached a remarkable role to this attribute (Luger and Goldstein 1991; Massey et al. 1992; Ferguson and Olofsson 2004), which has also been classified as a "key performance indicator" in Monck (2010), it seems to emerge that it is relatively unimportant (Shapley value of 0.08 in the global setting) vis a vis other dimensions of performance. However, this evidence can be partially explained moving to the analysis of the pairwise relationships between performance attributes characterizing the evaluation process. Before moving to the Interaction indexes, let us recall that Shapley values resulting from the aggregated game (column *global* in Table 6) are significantly different from the corresponding importance indexes arbitrarily assigned to the same attributes in Ferrara and Mavilia (2012), underlining the relevance of our approach in eliciting individual preferences.

Besides the importance of different attributes to determine the overall success of Science Parks, interactions between attributes disclose interesting results. As anticipated in the description of the methodology, simple weighted sum is not able to account for interactions among criteria and subsets of criteria. Table 7 presents the sign of the interactions between different attributes stemming from the analysis of the aggregate capacity. We recall that a positive interaction between two attributes means that the value assigned to high achievements on both attributes exceeds the sum of the values gained from the same achievements on two attribute separately. Conversely, a negative interaction between the two attributes i and j suggests the existence of some redundancy between them. First of all, let us report that at the aggregate level (specifically, the node at the top of the attributes' tree reported in Fig. 1) we find evidence of a complementarity between the two branches. It seems a quite natural finding: stakeholder are prone to reward SPs performing well in both the two areas. Now let us move to the more interesting analysis within each branch.

Table 7 reveals the existence of both positive and negative interactions. First, we observe different complementarities; focusing on the innovation branch, research centres

Table 7 Interactions for the aggregated game

Innovation	Research centers	Patents	Scientific network	Projects
Research centres	NA	+	−	+
Patents		NA	+	−
Scientific network			NA	+
Projects				NA

Entrepreneurship	Growth	Employees	Entropy	Geo-consistency
Growth	NA	−	−	+
Employees		NA	+	−
Entropy			NA	+
Geo-consistency				NA

are found to be complementary to patenting activity and projects, which in turns exhibit a positive interaction with the number of linkages with universities. Finally, scientific network is complementary to patents. From a decision making perspective, these results suggest that, for the aggregated decision maker, the SPs scoring high on both dimensions of each pairs should receive, *ceteris paribus*, a premium with respect to those scoring higher than previous ones but along only one of the two dimensions. An interpretation that is consistent with these evidences is that: according to our stakeholders, in a "well-functioning" SP the presence (and quantity) of research centres should ease the access of firms to the phases of R&D, which should improve the production of patents and projects that in turn supports the creation of scientific networks. In such a case, the decision maker would reward those SPs that are able to sustain all the steps of this cascade. Similarly looking at the entrepreneurship side, we find evidence of positive interactions between sectoral specialization (entropy) and both employment and proximity (geo-consistency) and between firms' growth and proximity. Some of these interactions might be interpreted straightforwardly. Since they are found, in general, between attributes referring to very different dimensions of performance (e.g sectoral specialization and jobs' creation or firms' patenting intensity and SP's connections with the academia), it could appear natural that respondents recognize a premium to those organizations which are able to perform well in both these different aspects. On the other side we observe also positive interactions between attributes that appear quite correlated (e.g. number of projects, which are often carried out by a network of organizations, and links with the academia, or research centres and patents); in those cases we also observe that respondents consider such attributes amongst the most important in general (Table 6) and therefore assign a premium to those SPs which successfully perform along all generally relevant dimensions.

On the contrary, negative interactions reveal the existence of some redundancy between attributes. We find evidence of this effect between patents and projects and, research centres and scientific network within the innovation side, and between firms' growth and both employees and sectoral specialization on the entrepreneurship side. This means that there is something like a negative premium from scoring highly on both of these pairs of dimensions rather than on other pairs. This effect might be due to the correlation between the two dimensions. For example, a park promoting a large number of projects (they are often research projects, Ferrara and Mavilia 2012) will probably host firms more prone to innovation, but the negative interaction suggests that, according to our respondents, the park should not be rewarded twice. The same explanation appears reasonable for the case of research centres and academic networks: parks with more research centres attract a larger number of university researchers and this leads to numerous connections between the park and the academia. Finally, it is relevant to discuss the negative interaction registered within the entrepreneurship branch between growth and employees. The group of our respondent is closely linked to academic research and it probably recognizes that sales growth is very often accompanied by job creation effects. For example, Delmar (2006) analyses a sample of small and medium sized firms reporting a correlation above 0.8 for sales and employment growth and Brouwer et al. (1993) is only one among several studies reporting sales' growth among the determinants of job creation. Hence, it appears reasonable that our respondents assign a large importance to firms' growth and find out a redundancy of the latter with the employee attribute (Tables 6 and 7).

Results II: comparative analysis of Italian Science Parks

The final step of our procedure consists in using the output of the elicitation exercise described in the previous sections to compare *actual* Italian SPs and find out those performing best and worse according to the aggregate preferences of our stakeholders.

Ranking Italian Science Parks

Our sample is composed as follows. All active Italian SPs in 2012 are included if they provide services to sustain firms' research activities, and at least one incubation structure is hosted within the park's premises. In case they act as virtual SPs, we require them to be associated to a minimum of one external research centre or University and to offer some services facilitating business activities of associated firms.[21] This procedure led to a final sample of 56 SPs, which geographically covers the whole territory of Italy with a stronger density in North-West regions. Information about performances of each SP has been retrieved from Ferrara and Mavilia (2012), Ferrara et al. (2012) and Ferrara et al. (2014). As previously introduced, the best scoring park has been assigned a value of 100, and all the others have received a score such that the percentage difference with respect to the immediately preceding park in the ranking is preserved.[22] In addition, such procedure guarantees that all dimensions are homogenized in terms of their units of measurement. Table 8 reports normalized performances of all Italian Science Parks in our sample along the eight dimensions described in "Dimensions and attributes" section. One of the emerging features is that the degree of heterogeneity in performances is extremely high. In particular, there is no SP which has been able to score high on all dimensions. For example consider AREA Science Park, which is the oldest and most known Italian SP. It has been often indicated as the reference point for all other SPs and an example of success (Battaglia et al. 2012; Liberati et al. 2015). It is the Italian structure with the largest presence of research centres and amongst the more specialized. However, the firms it hosts are poorly innovative (tend to exhibit few patent applications with respect to the leading parks in this dimension) and, relatively to the others, participate in few projects. Likewise AREA, also other best performing parks in some dimension present different weaknesses. On the other side, there is a set of SPs whose performances are relatively poor along all dimensions, but even within such group it is difficult to compare different structures. This evidence confirms the relevance of the approach we propose; in a context where performances are highly heterogeneous, a tool that identifies best performers on the basis of stakeholder preferences might be useful in a variety of decision making problems (e.g. how to distribute funds or where to locate a research unit).

The final aggregation of the multidimensional performances reported by Table 8 becomes now simple and requires as unique ingredient the aggregate capacity we have discussed above. The computation of the aggregate index to evaluate and compare performances of SPs boils down in a straightforward application of the formula for the

---

[21] These services might include: business, legal and marketing consultancy, facilitated access to credit, facilitated renting fares, participation in sponsored events and fairs.

[22] In the case of firms' growth, a value of 0 is assigned when the observed performance exhibits a negative value. Note that this is a conservative choice motivated by the willingness of preventing an excessive penalization of bad performers in a comparative evaluation

Table 8 Normalized observed performance of Italian Science Park (100 = max)

SP	Innovation				Entrepreneurship			
	Res. Centr.	Patents	Sci. Netw.	Proj.	Growth	Empl.	Geo- cons.	Entropy
Agemont	9.09	0	10	3.55	0	0.15	0	50.31
AGIRE	13.64	1.39	20	14.19	0	4.74	15.75	79.2
AGROBIOPESCA	36.36	0.35	30	6.76	0	2.38	29.54	68.03
AREA Science Park	100	0.36	20	6	24.78	2.39	22.87	64.58
Asteria	13.64	0.22	40	27.51	0.78	1.42	13.6	69.79
BicLiguria	0	0	0	1.16	10.53	21.05	26.15	79.36
Bioindustry Park	22.73	0.96	10	6.66	44.48	0.24	42.84	42.84
BiopMed	54.55	6.59	30	100	2.4	3.05	39.93	66.33
Campania Innovazione	22.73	2.01	30	3.37	14.91	0.47	21.98	52.98
Cittá Studi - Polo Pointex	13.64	4.94	10	9.46	13.22	20.42	8.89	68.52
ComoNEXT	13.64	0.81	100	27.03	45.47	0.14	11.27	81.26
Consorzio IBIS	4.55	100	10	12.16	8.99	16.74	12.98	52.62
Cosmob Spa	0	41.03	30	59.46	0	9.47	15.82	35.06
Digital Hub	4.55	1.16	20	3.12	12.57	24.47	20.51	64.75
DLTM	13.64	40.17	10	12.16	17.73	43.84	26.95	44.38
DTSMICRONANO	22.73	0	30	18.24	13.28	100	92.18	44.97
EASTGATE PARK	0	3.35	0	0	0	5.83	24.82	18.69
ENERMHY	4.55	12.23	30	32.43	7.23	18.05	28.4	89.48
Environment Park	18.18	0.13	30	11.44	6.15	13.89	26.25	71.63
Friuli Innovazione	13.64	0.11	10	12.79	0	0.14	11.33	42.44
IAM Innovazione Automotive	9.09	27.05	20	8.11	0	8.49	44.24	90.55
INSUBRIAS Biopark	0	3.35	10	7.43	20.04	24.64	15.82	0
Kilometro Rosso	13.64	1.57	20	3.55	8.07	4.36	9.54	18.69
Meccano spa	0	3.28	20	1.58	2.14	12.21	17.72	79.01
MESAP	40.91	37.87	20	31.08	0.02	19.44	17.46	87.55
Molise Innovazione	4.55	1.16	10	13.1	9.73	2.48	70.42	71.16
NAVTEC	13.64	7.11	30	0	39.59	80.89	54.04	49.07
Parco Innov.Valtellina	13.64	0	0	5.41	45.91	1.85	42.94	7.45
Parco Scientifico Romano	18.18	0.89	10	8.94	61.45	0.01	13.48	31.54
Parco Tecnologico Padano	4.55	0.64	10	27.75	44.07	0.15	19.67	17.16
Pepinieres	0	2.51	10	0	6.66	1.03	8.04	43.39
Polaris	31.82	2.23	20	13.06	100	0.22	61.84	37.43
POLIBRE	4.55	7.03	30	17.57	0.02	6.9	36.01	70.96
Polight	13.64	28.85	30	45.03	0.16	4.27	11.65	100
Polo ICT Abruzzo	4.55	0.73	10	6.08	0	4.5	27.05	65.01
Polo ICT Piemonte	45.45	38.43	40	26.11	18.04	49.01	19.79	56.65
Polo Innov. Pordenone	0	9.49	20	4.46	0	16.46	7.25	69.82

Table 8 continued

SP	Innovation				Entrepreneurship			
	Res. Centr.	Patents	Sci. Netw.	Proj.	Growth	Empl.	Geo-cons.	Entropy
Polo Innov. Lago Maggiore	9.09	2.43	30	18	0	3.39	18.6	68.84
Polo Scienze della Vita	13.64	1.77	50	6.59	5.18	5.84	42.81	25.6
Polo Tecnologico di Navacchio	18.18	0.2	50	3.84	67.45	0.34	5.93	44.73
Polo Ticass	13.64	2.34	10	52.7	9.36	3.88	7.87	68.42
Porto Conte Ricerche	18.18	0.17	20	4.77	10.42	0.12	98.85	41.53
Proplast	36.36	34.36	40	53.72	3.22	28.75	42.86	92.42
PST della Valle Scrivia	4.55	1.83	30	3.04	39.41	14.26	19.37	66.27
PST Salerno	0	0	20	15.5	17.14	1.13	100	25.35
PST Sicilia	4.55	0.74	30	13.71	0	93.4	38.92	64.52
Servitec	4.55	0.44	10	6.08	3.28	19.94	29.26	0
SI4LIFE	4.55	1.49	10	4.05	2.62	1.4	19.33	40.28
STARTCUBE	0	0.58	10	20.27	0	0.36	1.9	53.17
Tecnogranda	40.91	39.53	60	8.11	5.82	10.35	29.91	86.96
Tecnoparco Lago Maggiore	18.18	0	10	1.9	12.19	1.05	8.23	70.52
Tecnopolis	13.64	4.92	10	2.03	5.13	8.74	69.57	68.8
Tecnopolo Tiburtino	22.73	0.11	30	4.77	25.44	0.81	6.27	69.34
Trentino sviluppo	0	0.68	10	8.78	43.13	1.12	10.74	77.02
VEGA PARK PST	50	0.15	20	3.91	1.48	8.27	25.85	89.1
VMMP	4.55	0	20	0	37.19	0.32	67.02	28.03

Choquet integral expressed in "The structure of MAVT with Choquet integration" section where μ is now the aggregate capacity.[23]

Table 9 reports 20 out of 56 SPs listed in descending order according to the value of the associated Choquet integral, where the highest (resp. lowest) rank is assigned to the best (resp. worst) performing structure. Separate ranking for the two areas, i.e. entrepreneurship and innovation, are reported as well. Tables 10 and 11 in the "Appendix 2" show the values of the integral for all Italian SPs involved in the analysis.

It is worth to notice two features. First, the role played by interactions is fundamental. The SP resulting the best performer according to our stakeholders' preferences is relatively well scoring in each of the eight dimensions, but it never achieves the highest level of performances. Moreover, it does not lead neither the innovation nor the entrepreneurship area, but it finds amongst the top 10 performers in both (unique case together with ComoNEXT). Hence, it is the structure which gains the most from the complementarity identified between the two branches (see "Properties of the aggregation" section). This leads to another consideration: the difference with respect to results that could had been obtained with a simple average is large. If one should try to count the number of SPs that

---

[23] For sake of simplicity, we assume, for the aggregate decision maker, $u(x) = x$. However, the approach is robust to different specifications. Comparison tables are available from the authors.

Table 9 Choquet integral for actual science parks

	Innovation		Entrepreneurship		Final	
#	SP	Choquet	SP	Choquet	SP	Choquet
1	BiopMed	47.87	Polaris	58.23	Proplast	38.61
2	Proplast	40.78	NAVTEC	52.58	BiopMed	37.61
3	Polo ICT Piemonte	38.11	DTSMICRONANO	51.38	Polo ICT Piemonte	35.87
4	Tecnogranda	37.11	PST Sicilia	41.94	Tecnogranda	33.79
5	MESAP	33.59	ComoNEXT	39.42	ComoNEXT	32.94
6	AREA Science Park	32.90	Polo Tecnologico di Navacchio	37.76	Polaris	31.18
7	Consorzio IBIS	32.88	Trentino sviluppo	37.58	AREA Science Park	31.05
8	Cosmob Spa	31.27	PST della Valle Scrivia	37.58	MESAP	30.95
9	ComoNEXT	30.26	Proplast	37.21	DTSMICRONANO	28.92
10	Polight	28.62	Bioindustry Park	35.21	Consorzio IBIS	27.94
47	VMMP	5.16	INSUBRIAS Biopark	15.25	Polo ICT Abruzzo	10.64
48	Meccano spa	5.11	Pepinieres	14.68	Friuli Innovazione	10.07
49	Polo ICT Abruzzo	5.00	SI4LIFE	14.23	Kilometro Rosso	9.36
50	Servitec	4.93	Cosmob Spa	13.26	STARTCUBE	9.11
51	INSUBRIAS Biopark	4.75	STARTCUBE	13.25	INSUBRIAS Biopark	8.38
52	SI4LIFE	4.67	Agemont	12.19	SI4LIFE	7.97
53	Trentino sviluppo	4.39	Friuli Innovazione	12.15	Agemont	7.77
54	Pepinieres	2.60	Kilometro Rosso	10.17	Pepinieres	6.83
55	EASTGATE PARK	0.89	EASTGATE PARK	9.88	Servitec	6.52
56	BicLiguria	0.27	Servitec	9.69	EASTGATE PARK	4.06

are assigned a different position under the two aggregation models, she will discover that more than 80 % of SPs would do.

Discussion, limitations and future developments

As we have seen, recently, the issue of enhancing and promoting innovation activity of both public and private institutions has become extremely attractive for academics as well as policy-makers. Furthermore, we have clearly pointed out throughout the paper the role that SPs can play. Different groups of stakeholders might be interested in comparing such structures and they might need a flexible tool to support their search for best and worse performer that relies on their specific preferences. However evaluating SPs is a complex exercise, especially when performances are strongly heterogeneous; our approach is a novelty in the literature and the empirical application may be viewed as a first attempt to consistently evaluate SPs, taking into account their multidimensional nature and the

subjective preferences of parks' stakeholders. Obviously, this study can be easily extended in several directions and, therefore, might be seen as a starting point for future research.

In particular, the small sample does not take into consideration all possible actors involved in a SP. Hence, a larger sample including, for instance, managers of the parks as well as managers of firms hosted and representatives of local governments can surely strengthen the mechanism through which we derive the weights of each attribute. This would not affect the methodological approach but can make empirical applications more robust. For example, we might expect a difference in the importance index assigned to job-creation had we included park managers and, especially, policy makers, who are traditionally more concerned with this aspect that academic researchers might be.

Moreover, we focused on Italian SPs but our approach it is completely replicable in order to evaluate SPs in several other countries and eventually compare the results. It would be interesting to assess whether there are differences in the relative importance of dimensions or attributes as well as in the interactions among them within a different setting.

Finally, we have considered some specific dimensions and performance attributes in order to rank Italian SPs, although, broadening the set of dimensions can result in a more precise elicitation of preferences. With respect to this point, we underline that a narrower segmentation between different organizations (e.g. Research, Science and University Parks) might provide might allow to identify best and worst performers within each specific category and provide more detailed information to, for example, policy makers.

Conclusions

Evaluating and comparing Science Parks' performances are still open problems both for the academic literature, practitioners and policy makers. In this paper we develop and apply a methodology that allows us to non-arbitrarily rank SPs, along sets of attributes that can be organized hierarchically and account explicitly for interactions (complementarities and redundancies) among attributes and sets of attributes. In addition, our approach allows for the participation of any number of experts, whose preferences are rapidly elicited through a questionnaire, which is included in the supplementary material. This guarantees an advantage for practical applications. In a second stage we have then applied the described methodology to the case of Italian Science Parks. This constitutes a pilot study in the field and, as a consequence, relies on a small number of parks' stakeholders. Despite this limitation, we have found that any linear evaluating function (that is, any weighted average of different attributes of performance) turns out to be inadequate for the purpose of comparing multi-dimensional and complex organizations as Science Parks. This is the case because interactions among attributes play a significant role. In particular, we found that innovation related performances are deemed more relevant than those linked to entrepreneurship. What matters to the group of our stakeholders is SPs' effectiveness in inducing innovation and supporting research related activities within firms hosted in the park, while their role as producers of "economic value" (e.g. job-positions) is relatively less important. Almost all research related attributes of performance are found to be complementary: a well-functioning park should be able to sustain them all. On the other side, our stakeholders recognize a redundancy between two of the most used performance measures in the literature (sales' growth and job-creation). Using all these information to rank actual Italian Science Parks we find evidence of a relevant difference between our

model and a simple average. Best performers are those parks able to score relatively high on key performance dimensions, even though they are never leading in a singular dimension. Further research will focus on the inclusion of a broader set of stakeholders within the evaluation procedure, on the analysis of differences between such sets and on a continuous monitoring and updating of Science Park' s activities, in the hope to give access to a useful signal both to potential entrepreneurs, parks' managers and policy makers.

Appendix 1

In this Appendix, we present in greater details the technical features of the methodology involved in the evaluation of SPs. As we have seen, the application of MAVT based on the Choquet integral requires the identification of a capacity. In what follows, we firstly describe an alternative representation of the the Choquet integral, then we will formally specify the definition of the two indexes of aggregation which have been used in the analysis.

Any set function $\mu : \mathcal{P}(N) \rightarrow \mathbb{R}$ can be uniquely expressed in terms of its Möbius transform by:

$$\mu(T) = \sum_{S \subseteq T} m_\mu(S), \ \forall T \subseteq N$$

where the set function $m_\mu : \mathcal{P}(N) \rightarrow \mathbb{R}$ is called Möbius transform of a capacity μ and it is defined by:

$$m_\mu(S) = \sum_{T \subseteq S} (-1)^{s-t} \mu(T), \ \forall S \subseteq N$$

Now we can rewrite the Choquet integral in terms of the Möbius representation of a capacity. For any $u(x) := (u_1(x_1), \ldots, u_n(x_n)) \in \mathbb{R}$, the Choquet intergral of x w.r.t μ is given by:

$$C_{m_\mu}(u(x)) = \sum_{T \subseteq N} m_\mu(T) \bigwedge_{i \in T} u_i(x_i)$$

where $\bigwedge$ represents the minimum operator. The notation C_{m_μ} clarifies the fact that the Choquet integral is computed w.r.t. the Möbius transform of the capacity μ.

As we have mentioned before, the concept of *k-additivity* can capture the trade-off between the complexity of the capacity and its modeling ability. A capacity μ on N, indeed, is totally defined by a reasonably large number of coefficients, i.e. $2^n - 2$, and therefore such complexity ma result prohibitive in some applications.

Definition Let $k \in 1, \ldots, n$. A capacity μ on N is said to be *k-additive* if its Möbius representation satisfies $m_\mu(T) = 0 \ \forall T \subseteq N$ such that $t \geq k$ and there exists at least one subset T of cardinality k such that $m_\mu(T) \neq 0$.

Clearly, the notion of 1-additivity coincides with that of additivity. Moreover, a capacity that is *k-additive* ($k < n$) turns out to be completely defined by the knowledge of $\sum_{l=1}^{k} \binom{n}{l}$.

In order to clarify the behaviour of the Choquet integral as an aggregation operator, we relied on two different measures:

- The Importance Index
- The Interaction Index

Here, we will define them more formally.

As we have seen, the overall importance of an attribute $i \in N$ can be captured by means of its Shapley value which is formally defined by:

$$\phi(\mu, i) := \sum_{T \subseteq N \setminus i} \frac{(n - t - 1)! t!}{n!} [\mu(T \cup i) - \mu(T)]$$

where for each subset of attribute $S \subseteq N$, $\mu(S)$ can be interpreted as the importance os S in the decision problem. Therefore, the Shapley value can be defined as a weighted average value of the marginal contribution $\mu(T \cup i) - \mu(T)$ of element i alone in all combinations.

Moreover, in terms of its Möbius transform the Shapley value takes a really simple form:

$$\phi(\mu, i) = \sum_{T \ni i} \frac{1}{t} m_\mu(T)$$

For what concerns the interaction index, instead, following Murofushi and Soneda (1993), we can consider the interaction index for i and j as the average value of this marginal interaction. Therefore, setting

$$(\Delta_{ij}\mu) := \mu(T \cup ij) - \mu(T \cup i) - \mu(T \cup j) + \mu(T)$$

we have that the interaction index of attributes i and j related to μ is given by:

$$I(\mu, ij) := \sum_{T \subseteq N \setminus ij} (\Delta_{ij}\mu)(T)$$

Clearly, such index takes a negative value as long as i and j are positively correlated and, as a consequence, a positive value when i and j are negative correlated. Furthermore, $I(\mu, ij) \in [-1, 1]$ $\forall i, j \subseteq N$. The value 1(respectively, -1) denotes maximum complementarity (substitutivity) between attributes i and j.

Appendix 2

See Tables 10 and 11.

Table 10 Choquet integral for all actual science parks

#	Innovation		Entrepreneurship	
	SP	Choquet	SP	Choquet
1	BiopMed	47.87	Polaris	58.23
2	Proplast	40.78	NAVTEC	52.58
3	Polo ICT Piemonte	38.11	DTSMICRONANO	51.38
4	Tecnogranda	37.11	PST Sicilia	41.94
5	MESAP	33.59	ComoNEXT	39.42
6	AREA Science Park	32.90	Polo Tecnologico di Navacchio	37.76
7	Consorzio IBIS	32.88	Trentino sviluppo	37.58
8	Cosmob Spa	31.27	PST della Valle Scrivia	37.58
9	ComoNEXT	30.26	Proplast	37.21
10	Polight	28.62	Bioindustry Park	35.21
11	DLTM	19.64	Polo ICT Piemonte	34.22
12	Polo Ticass	19.02	Parco Scientifico Romano	33.31
13	VEGA PARK PST	18.87	Molise Innovazione	33.14
14	Asteria	18.82	ENERMHY	33.10
15	ENERMHY	18.70	VMMP	32.54
16	AGROBIOPESCA	18.16	BicLiguria	32.18
17	DTSMICRONANO	17.10	Tecnopolis	31.93
18	Polaris	16.84	DLTM	31.38
19	Polo Tecnologico di Navacchio	15.95	IAM Innovazione Automotive	31.12
20	IAM Innovazione Automotive	15.88	Tecnogranda	30.44
21	Polo Scienze della Vita	15.71	Porto Conte Ricerche	29.94
22	Environment Park	14.05	AREA Science Park	29.75
23	Campania Innovazione	13.86	Digital Hub	29.17
24	Tecnopolo Tiburtino	13.72	PST Salerno	28.95
25	Poloovazione Lago Maggiore	13.69	MESAP	28.49
26	POLIBRE	13.48	VEGA PARK PST	28.24
27	NAVTEC	11.88	Tecnopolo Tiburtino	27.96
28	AGIRE	11.83	Parco Innov. Valtellina	27.76
29	PST Sicilia	10.81	Citt Studi Polo Pointex	27.46
30	Porto Conte Ricerche	10.47	Polight	27.13
31	Bioindustry Park	10.24	Environment Park	27.08
32	Parco Tecnologico Padano	10.10	Meccano spa	25.61
33	Parco Scientifico Romano	9.58	POLIBRE	24.67
34	Citt Studi Polo Pointex	9.55	Parco Tecnologico Padano	24.28
35	Friuli Innovazione	9.20	BiopMed	24.17
36	Kilometro Rosso	9.18	Tecnoparco del Lago Maggiore	23.45
37	PST della Valle Scrivia	8.31	AGIRE	22.83
38	PST Salerno	7.87	Campania Innovazione	22.48
39	Tecnopolis	7.76	Polo Ticass	22.32
40	Tecnoparco Lago Maggiore	7.56	Consorzio IBIS	21.91
41	Poloovazione di Pordenone	7.56	AGROBIOPESCA	21.84
42	STARTCUBE	7.05	Polo Innov. di Pordenone	21.68

Table 10 continued

#	Innovation SP	Choquet	Entrepreneurship SP	Choquet
43	Molise Innovazione	6.86	Polo ICT Abruzzo	21.20
44	Digital Hub	6.25	Polo Innov. Lago Maggiore	20.48
45	Agemont	5.51	Asteria	19.72
46	Parco Innov. della Valtellina	5.17	Polo Scienze della Vita	16.29
47	VMMP	5.16	INSUBRIAS Biopark	15.25
48	Meccano spa	5.11	Pepinieres	14.68
49	Polo ICT Abruzzo	5.00	SI4LIFE	14.23
50	Servitec	4.93	Cosmob Spa	13.26
51	INSUBRIAS Biopark	4.75	STARTCUBE	13.25
52	SI4LIFE	4.67	Agemont	12.19
53	Trentino sviluppo	4.39	Friuli Innovazione	12.15
54	Pepinieres	2.60	Kilometro Rosso	10.17
55	EASTGATE PARK	0.89	EASTGATE PARK	9.88
56	BicLiguria	0.27	Servitec	9.69

Table 11 Choquet integral for all actual science parks

#	Final index SP	Choquet
1	Proplast	38.61
2	BiopMed	37.61
3	Polo ICT Piemonte	35.87
4	Tecnogranda	33.79
5	ComoNEXT	32.94
6	Polaris	31.18
7	AREA Science Park	31.05
8	MESAP	30.95
9	DTSMICRONANO	28.92
10	Consorzio IBIS	27.94
11	Polight	27.50
12	NAVTEC	26.07
13	Cosmob Spa	23.58
14	ENERMHY	23.45
15	DLTM	23.44
16	Polo Tecnologico di Navacchio	23.38
17	VEGA PARK PST	21.83
18	PST Sicilia	21.63
19	IAM Innovazione Automotive	20.98
20	Polo Ticass	19.84
21	AGROBIOPESCA	19.13
22	Bioindustry Park	18.89
23	Asteria	18.79

Table 11 continued

	Final index	
#	SP	Choquet
24	PST della Valle Scrivia	18.52
25	Tecnopolo Tiburtino	18.50
26	Environment Park	18.40
27	Parco Scientifico Romano	17.80
28	POLIBRE	17.19
29	Porto Conte Ricerche	17.17
30	Campania Innovazione	16.66
31	Tecnopolis	16.18
32	Trentino sviluppo	16.06
33	Molise Innovazione	16.04
34	Polo Innov. Lago Maggiore	15.84
35	Citt Studi Polo Pointex	15.71
36	Polo Scienze della Vita	15.62
37	AGIRE	15.51
38	PST Salerno	15.19
39	Parco Tecnologico Padano	14.93
40	VMMP	14.76
41	Digital Hub	14.25
42	Parco Innov. della Valtellina	13.07
43	Tecnoparco Lago Maggiore	13.05
44	Polo Innov. Pordenone	12.42
45	Meccano spa	12.28
46	BicLiguria	11.56
47	Polo ICT Abruzzo	10.64
48	Friuli Innovazione	10.07
49	Kilometro Rosso	9.36
50	STARTCUBE	9.11
51	INSUBRIAS Biopark	8.38
52	SI4LIFE	7.97
53	Agemont	7.77
54	Pepinieres	6.83
55	Servitec	6.52
56	EASTGATE PARK	4.06

References

ANGLE (2003). *Evaluation of the past and future economic contribution of the UK science park movement.* United Kingdom Science Park Association.

Battaglia, R., Lamperti, F., & Siligato, L. (2012). AREA SP: potere alleccellenza. In M. Ferrara & R. Mavilia (Eds.), *Dai distretti tecnologici ai poli di innovazione*. Milan: Egea.

Bernoulli, J. (1713). *Ars conjectandi, opus posthumum. Accedit Tractatus de seriebus infinitis, et epistola gallic scripta de ludo pilae reticularis*. Basel: Thurneysen Brothers.

Brouwer, E., Kleinknecht, A., & Reijnen, J. O. (1993). Employment growth and innovation at the firm level. *Journal of Evolutionary Economics*, 3(2), 153–159.

Choquet, G. (1953). Theory of capacities. *Annales de l'Institut Furier*, 5, 131–295.

Colombo, M. G., & Delmastro, M. (2002). How effective are technology incubators? Evidence from italy. *Research policy, 31*(7), 1103–1122.

Dabrowska, J. (2011). Measuring the success of science parks: performance monitoring and evaluation. In *Paper for the XXVIII World Conference on Science and Technology Parks. IASP*.

Debreu, G. (1960). Topological methods in cardinal utility theory. In *Mathematical Methods in the Social Sciences*. Stanford University Press.

Delmar, F. (2006). Measuring growth: methodological considerations and empirical results. In *Entrepreneurship and the Growth of Firms* (Vol. 1, pp. 62–84). UK: Edward Elgar Publishing Cheltenham.

Felsenstein, D. (1994). University-related science parks, seedbeds or enclaves of innovation. *Technovation, 14*, 93–110.

Ferguson, R., & Olofsson, C. (2004). Science parks and the development of ntbfslocation, survival and growth. *The Journal of Technology Transfer, 29*(1), 5–17.

Ferrara, M., & Mavilia, R. (2012). *Dai distretti industriali ai poli di innovazione: L'Italia nel Mediterraneo*. Milan: Egea.

Ferrara, M., Lamperti, F., & Mavilia, R. (2014). The impact of technopoles and science parks on the regional and local innovation systems. In *Design a Pattern of Sustainable Growth. Innovation, Education, Energy and Environmnent*. Asers Publishing.

Ferrara, M., Mavilia, R., & Lamperti, F. (2012). The effects of innovation poles and science parks on regional economies in italy. In *Proceedings of the 3rd international conference on communication and management in technological innovation and academy* (pp. 196–202).

Fish, M. S. (2006). Stronger legislatures, stronger democracies. *Journal of Democracy, 17*(1), 5–20.

Gerardi, D., McLean, R., & Postlewaite, A. (2009). Aggregation of expert opinions. *Games and Economic Behavior, 65*(2), 339–371.

Grabisch, M. (1996). The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research, 89*(3), 445–456.

Grabisch, M., Kojadinovic, I., & Meyer, P. (2008). A review of methods for capacity identification in choquet integral based multi-attribute utility theory: Applications of the kappalab r package. *European Journal of Operational Research, 186*(2), 766–785.

Guy, I. (1996). A look at aston science park. *Technovation, 16*(5), 217–218.

Hansson, F., Husted, K., & Vestergaard, J. (2005). Second generation science parks: From structural holes jockeys to social capital catalysts of the knowledge society. *Technovation, 25*(9), 1039–1049.

Hodgson, B. (1996). A methodological framework to analyse the impact of science and technology parks. In M. Guedes & P. Formica (Eds.), *The Economics of Science Parks*. : IASP and AURRP.

Hsu, H.-M., & Chen, C.-T. (1996). Aggregation of fuzzy opinions under group decision making. *Fuzzy Sets and Systems, 79*(3), 279–285.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society, 47*, 263–291.

Krantz, D., Luce, D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement, Vol. I: Additive and polynomial representations*. : Dover Books on Mathematics.

Labreuche, C., & Grabisch, M. (2003). The choquet integral for the aggregation of interval scales in multicriteria decision making. *Fuzzy Sets and Systems, 137*, 11–16.

Lamperti, F., Mavilia, R., & Castellini, S. (2015). The role of science parks: A puzzle of growth, innovation and R&D investments. *The Journal of Technology Transfer*, 1–26.

Liberati, D., Marinucci, M., & Tanzi, G. (2015). Science and technology parks in italy: main features and analysis of their effects on the firms hosted. *The Journal of Technology Transfer*, pp 1–36.

Link, A. N., & Link, K. R. (2003). On the growth of us science parks. *The Journal of Technology Transfer, 28*(1), 81–85.

Link, A. N., & Scott, J. T. (2003). Us science parks: The diffusion of an innovation and its effects on the academic missions of universities. *International Journal of Industrial Organization, 21*(9), 1323–1356.

Löfsten, H., & Lindelöf, P. (2002). Science Parks and the growth of new technology-based firms-academic-industry links, innovation and markets. *Research Policy, 31*(6), 859–876.

Luger, M. I., & Goldstein, H. A. (1991). *Technology in the garden: Research parks and regional economic development*. Chapel Hill: Univ of North Carolina Press.

Marichal, J.-L. (2000). Behavioral analysis of aggregation in multicriteria decision aid. In *Preferences and decisions under incomplete knowledge* (pp. 153–178). Springer.

Massey, D., Quintas, P., & Wield, D. (1992). *High-tech fantasies. Science parks in society, science and space*. London: Routledge.

Meyer, P., & Ponthière, G. (2011). Eliciting preferences on multiattribute societies with a choquet integral. *Computational Economics, 37*(2), 133–168.

Minguillo, D., & Thelwall, M. (2015). Which are the best innovation support infrastructures for universities? Evidence from R&D output and commercial activities. *Scientometrics*, *102*(1), 1057–1081.

Minguillo, D., Tijssen, R., & Thelwall, M. (2015). Do science parks promote research and technology? A scientometric analysis of the UK. *Scientometrics*, *102*(1), 701–725.

Monck, C. (2010). Performance monitoring and evaluation. In *UKSPA conference proceedings*.

Monck, C., & Peters, K. (2009). Science parks as an instrument of regional competitiveness: Measuring success and impact. In *IASP 2009 annual conference proceedings*.

Moon, J. H., & Kang, C. S. (1999). Use of fuzzy set theory in the aggregation of expert judgments. *Annals of Nuclear Energy*, *26*(6), 461–469.

Murofushi, T., & Soneda, S. (1993). Techniques for reading fuzzy measures (iii): Interaction index. In *Proocedings of the 9th Fuzzy System Symposium*, Sapporo (Japan).

Pinar, M., Cruciani, C., Giove, S., & Sostero, M. (2014). Constructing the feem sustainability index: A choquet integral application. *Ecological Indicators*, *39*, 189–202.

Reinhart, C. M., & Rogoff, K. S. (2014). Recovery from financial crises: Evidence from 100 episodes. *American Economic Review*, *104*(5), 50–55.

Saublens, C. et al. (2007). Regional research intensive clusters and science parks. *European Commission, ec. europa. eu/research/regions/pdf/sc_park. pdf*.

Shapley, L. S. (1953). Stochastic games. *Proceedings of the National Academy of Sciences of the United States of America*, *39*(10), 1095.

Siegel, D. S., Westhead, P., & Wright, M. (2003). Assessing the impact of university science parks on research productivity: Exploratory firm-level evidence from the united kingdom. *International Journal of Industrial Organization*, *21*(9), 1357–1369.

Smith, J. H. (1973). Aggregation of preferences with variable electorate. *Econometrica: Journal of the Econometric Society*, pp 1027–1041.

Squicciarini, M. (2009). Science parks, knowledge spillovers, and firms innovative performance. evidence from finland. Economics Discussion Paper 2009-32, Kiel Institute for the World Economy.

Sugeno, M. (1974). Theory of fuzzy integrals and its applications. Ph.D. Thesis, Tokio Institute of Technology.

Vedovello, C. (1997). Science parks and university-industry interaction: Geographical proximity between the agents as a driving force. *Technovation*, *17*(9), 491–531.